JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 57, 20–26 (1977)

On the Robustness of Gaussian Detection

A. F. GUALTIEROTTI*

Département de Mathématiques, École Polytechnique Fédérale, CH 1007 Lausanne, Switzerland

Submitted by Gian-Carlo Rota

When trying to detect a Gaussian L_2-signal imbedded in Gaussian L_2-noise, one has to consider the equivalence of the corresponding induced measures. Since equivalence is a rather "unstable" condition and since in practice only finite dimensional distributions (marginals) are available, it is of interest to know how inference is affected by the failure to know "exactly" the induced distributions. Here we show that, if the weak covariance operators at hand behave adequately, the fact that they are not trace-class does not prevent some form of detection.

1. INTRODUCTION

In this paper we use a characterization of the norms of Hilbert spaces that permits extensions of Gaussian cylinder set measures to probability measures [1] to investigate a problem in statistical communication theory, namely, the nonsingularity of detection in the Gaussian case. It is known [2] that detection is always either singular or nonsingular. However, since the parameters (mean and variance) are never exactly known and since nonsingularity is a very unstable condition (for example, in the Gaussian case, the two covariances C and t^2C correspond to a singular problem, as soon as t is not 1), it is important to investigate how detection is affected, when the parameters are not exactly known.

Here we consider the case when covariance operators are actually weak covariances, i.e., bounded, linear, positive, and self-adjoint, but not trace-class. We will show that nonsingularity holds in a certain sense, provided weak covariances behave, with respect to each other, as if they were "bona fide" convariances, and we give the corresponding likelihood ratios.

* Research supported by the Office of Naval Research under Contract NR 00014-67-A-0321-0006.

ISSN 0022-247X

2. PRELIMINARIES

H is a real and separable Hilbert space with norm $\|\cdot\|$ and inner product $(\cdot, \cdot)$. C denotes a linear, bounded, self-adjoint, and nonnegative operator and is called a weak covariance (operator). C always determines a Gaussian cylinder set measure as follows. A cylinder set is a set A in H of the form (h in $H/((h, h_1),..., (h, h_n))$ in B), where $h_1, ..., h_n$ are fixed in H and B is a Borel set of R^n. m_C, the cylinder set measure determined by C, is a function, additive on all the sets of type A, which, when restricted to h_i's in a finite dimensional subspace of H, is a probability measure. Furthermore, the value of m_C at A is given by a Gaussian density

$$m_C(A) = \int_A f(x)\, dx,$$

where f is the normal density on R^n, with mean zero and covariance matrix $((Ch_i, h_j))_{i,j=1}^n$.

In [1] the following facts are established. It is always possible to imbed H linearly, continuously, and densely into some real and separable Hilbert space K (E = imbedding map, $[\![\cdot]\!]$ is the norm of K, $[\cdot, \cdot]$ is the inner product of K) in such a way that $m_C' = m_C \cdot E^{-1}$ extends to a Gaussian probability measure M_C on the Borel sets of K, with covariance $\bar{C}$. A necessary and sufficient condition is that $[\![Eh]\!] = \| Th \|$, where T is a bounded linear extension of $SC^{-(1/2)}$ to H, S Hilbert-Schmidt and self-adjoint. We will need (from [1] also) that $\bar{C}^{1/2} = EC^{1/2}W$, $W: K \to H$ unitary, and that $E'E = T'T$.

Now suppose that we are confronted with a Gaussian detection problem and that, unknown to us, we deal with a weak covariance operator instead of a "bona fide" covariance. The question is : Do the ordinary procedures, at least theoretically, still carry through? We want to show that in a certain sense the answer can be yes.

Let C_1 and C_2 be the two weak injective covariances in question. We suppose that the procedure described above can be carried out simultaneously for C_1 and C_2 on the same K. This is the main restriction, for, if T_i, S_i, $i = 1, 2$, are the operators involved, we must have $T_1'T_1 = T_2'T_2$ ($\hat{T}$ in what follows), or equivalently, $C_1^{-(1/2)}S_1 = C_2^{-(1/2)}S_2 V$, V unitary; that is, we must be able to solve an operator equation, where S_i, $i = 1, 2$, are under our control. Assuming thus that we can handle this problem, we can then decide that our actual problem is nonsingular provided the problem on K is nonsingular. But this is useful only if the latter can be decided by looking at C_1 and C_2 only, and if we can then produce a likelihood ratio in terms of C_1 and C_2 also.

3. Equivalence of Extended Measures

Proposition 1. *Within the framework described in Section 2, the following two conditions are equivalent.*

(a) $\bar{C}_1 = \bar{C}_2^{1/2}(\bar{I} + \bar{B})\,\bar{C}_2^{1/2}$, *where* $\bar{I}$ *is the identity operator on* K, $\bar{B}$ *is Hilbert–Schmidt (trace-class) on* K, *and* -1 *does not belong to the eigenvalues of* $\bar{B}$.

(b) $C_1 = C_2^{1/2}(I + B)\,C_2^{1/2}$, *where* I *is the identity operator on* H, B *is Hilbert–Schmidt (trace-class) on* H, *and* -1 *does not belong to the eigenvalues of* B.

Proof. "(b) implies (a)" is quite similar to "(a) implies (b)" and thus we consider only the latter.

Knowing that $\bar{C}_2^{1/2} = EC_2^{1/2}W_2$, $W_2 \colon K \to H$ is unitary, and that $\bar{C}_1 = EC_1E'$, one sees that the equality of (a) can be written

$$EC_1E' = EC_2^{1/2}(I + W_2\bar{B}W_2')\,C_2^{1/2}E'.$$

But E is one-to-one and E' has dense range, so that the latter equality can be simplified to

$$C_1 = C_2^{1/2}(I + W_2\bar{B}W_2')\,C_2^{1/2}.$$

Now, if (e_n) is a complete orthonormal set in H, $(W_2'e_n)$ is a complete orthonormal set in K, and the equalities

$$\sum_n \| W_2\bar{B}W_2'e_n \|^2 = \sum_n [\![\bar{B}W_2'e_n]\!]^2,$$

$$\sum_n (W_2\bar{B}W_2'e_n\,,\, e_n) = \sum_n [\bar{B}W_2'e_n\,,\, W_n'e_n],$$

$$W_2\bar{B}W_2'h = -h,$$

$$\bar{B}W_2'h = -W_2'h,$$

give the result. ∎

Corollary. *Detection is nonsingular (in the adopted sense), provided* (b) *holds.*

Proof. (a) is a well-known necessary and sufficient condition (see [2], for example). ∎

4. Likelihood Ratios

When the detection problem is nonsingular, the optimum operation on the data requires that the Radon–Nikodym derivative be evaluated. We indicate below how the standard results apply in the present context.

Lemma 1.

(a) *$\bar{C}_2$ and $\hat{T}^{1/2}C_2\hat{T}^{1/2}$ have the same eigenvalues;*

(b) *let U be the unitary operator in the polar decomposition $E = U(E'E)^{1/2}$. Then*

(i) *$\bar{C}_2 f_n = a_n f_n$ implies $\hat{T}^{1/2}C_2\hat{T}^{1/2}U'f_n = a_n U'f_n$;*

(ii) *$\hat{T}^{1/2}C_2\hat{T}^{1/2}e_n = b_n e_n$ implies $\bar{C}_2 Ue_n = b_n Ue_n$;*

(c) *$\hat{T}^{1/2}C_2\hat{T}^{1/2}$ is trace-class.*

Proof. $a_n f_n = \bar{C}_2 f_n = EC_2E'f_n = U\hat{T}^{1/2}C_2\hat{T}^{1/2}U'f_n$; $b_n Ue_n = U\hat{T}^{1/2}C_2\hat{T}^{1/2}U'Ue_n = \bar{C}_2 Ue_n$. ▮

Proposition 2. *Suppose that, within the given framework, $C_1 = C_2^{1/2}(I + B)\, C_2^{1/2}$. Then*

(a) *if B is Hilbert–Schmidt and -1 is not an eigenvalues of B, one has $(dM_{C_1}/dM_{C_2})\,(k) = \lim_n L_n(k)$ a.e. k, where*

(i) $L_n(k) = -\frac{1}{2}\sum_{p=1}^{n}((s_p^{-1} - 1)\, Y_p^2(k) + \ln s_p)$,

(ii) *s_p is the* pth *eigenvalue of $I + B$,*

(iii) *$Y_p(k) = \sum_q a_q^{(p)} X_q(k)$, where $a_q^{(p)}$ is obtained from the eigenvectors e_n of $\hat{T}^{1/2}C_2\hat{T}^{1/2}$ and the eigenvectors g_n of $I + B$ by writing $g_p = \sum_q a_q^{(p)} e_q$ and $X_q(k) = b_q^{-1/2}[k, W_2'e_q]$, with $\hat{T}^{1/2}C_2\hat{T}^{1/2}e_q = b_q e_q$;*

(b) *if B is trace-class and -1 is not among its eigenvalues, one has*

$$\frac{dM_{C_1}}{dM_{C_2}}(k) = -\frac{1}{2}\sum_{n=1}^{\infty}(s_n^{-1} - 1)\, Y_n^2(k) - \frac{1}{2}\sum_{n=1}^{\infty}\ln s_n\,, \qquad \text{a.e.} \quad k.$$

Proof. The result is a consequence of [2] completed by the following remarks.

(i) The eigenvalues of $\bar{C}_2$, and its eigenvectors, are obtained with the help of Lemma 1,

(ii) the eigenvalues and eigenvectors of $\bar{I} + W_2'BW_2$ are obtained from those of $I + B$, since the equality $(I + B)\, g_n = s_n g_n$ gives

$$0 = [W_2'(I + B)\, W_2(W_2'g_n) - s_n(W_2'g_n), W_2'h].$$

Further, one uses $\bar{C}_1 = \bar{C}_2^{1/2}(\bar{I} + W_2'BW_2)\,\bar{C}_2^{1/2}$. ▮

We now consider conditions under which the likelihood ratio is a quadratic form. There are two necessary and sufficient conditions [2]. The first relates to the ranges of the covariance operators under consideration, the other is concerned with the existence of a bounded extension of a certain operator. The latter can be dealt with without additional hypotheses: we thus state Propositions 4 and 5 with the condition range$(\bar{C}_1)$ = range$(\bar{C}_2)$. We then investigate what additional assumptions on C_1 and C_2 are necessary to insure that range$(\bar{C}_1)$ = range$(\bar{C}_2)$. This is the content of Propositions 6 and 7.

PROPOSITION 3. *Within the given framework, $(\bar{C}_1^{-1} - \bar{C}_2^{-1})\,\bar{C}_2^{-1}$ has a bounded, Hilbert–Schmidt extension to K if and only if $\hat{T}^{-(1/2)}(C_1^{-1} - C_2^{-1})\,C_2^{1/2}$ has a bounded, Hilbert–Schmidt extension to H.*

Proof. We have $(\bar{C}_1^{-1} - \bar{C}_2^{-1})\,\bar{C}_2^{1/2} = U'\hat{T}^{-(1/2)}(C_1^{-1} - C_2^{-1})\,C_2^{1/2}W'_2$ and $W'_2\bar{C}_2^{1/2}k = S_2W'_2k$, so that $\|(\bar{C}_1^{-1} - \bar{C}_2^{-1})\,\bar{C}_2^{1/2}(\bar{C}_2^{1/2}k)\| \leq A\|\bar{C}_2^{1/2}k\|$ if and only if $\|\hat{T}^{-1/2}(C_1^{-1} - C_2^{-1})\,C_2^{1/2}(S_2W'_2k)\| \leq A\,\|S_2W'_2k\|$, which gives the simultaneous extension.

Now let (e_n) be a complete orthonormal set for K contained in range$(\bar{C}_2^{1/2})$. Then $W'_2e_n = S_2W'_2f_n$, for some f_n, and range$(S_2) \subseteq$ range$(C_2^{1/2})$ gives that (W'_2e_n) is a complete orthonormal set in range$(C_2^{1/2})$. Consequently the result follows from the equality

$$\|(\bar{C}_1^{-1} - \bar{C}_2^{-1})\,\bar{C}_2^{1/2}e_n\|^2 = \|\hat{T}^{1/2}(C_1^{-1} - C_2^{-1})\,C_2^{1/2}(W'_2e_n)\|^2.$$

PROPOSITION 4. *Suppose that, within the given framework, $M_{C_1} \equiv M_{C_2}$. Then $L = \ln(dM_{C_1}/dM_{C_2})$ is a quadratic form if and only if*

(i) range$(\bar{C}_1)$ = range$(\bar{C}_2)$,

(ii) $\hat{T}^{-(1/2)}(C_1^{-1} - C_2^{-1})\,C_2^{1/2}$ *has a Hilbert–Schmidt extension to H.*

Proof. Immediate consequence of [2] and Proposition 3. ∎

LEMMA 2. *Let $U: H \to K$ be unitary, $\bar{A}: K \to K$ and $A: H \to H$ be linear. Suppose that U*(domain(A)) = domain$(\bar{A})$ and that $\bar{A} = UAU'$. *Then*

(a) *$\bar{A}$ admits a closure if and only if A admits a closure,*

(b) *$\bar{A}$ is closed if and only if A is closed,*

(c) *$\bar{A}$ is symmetric if and only if A is symmetric.*

Proof. It is easy to check that the appropriate definitions are satisfied. ∎

PROPOSITION 5. *Assume that, within the given framework, $M_{C_1} \equiv M_{C_2}$, and that $L = \ln(dM_{C_1}/dM_{C_2})$ is a quadratic form, then*

$$L(k) = \tfrac{1}{2}[\bar{A}_0k, k] - \tfrac{1}{2}\ln|\bar{S}|,$$

where

(i) $\bar{A}_c = UA_cU'$

(ii) A_c *is the closure of* $\hat{T}^{-1/2}(C_1^{-1} - C_2^{-1})\,\hat{T}^{-1/2}$,

(iii) $\bar{S} = \bar{I} + W_2'SW_2$, *with* S *the Hilbert–Schmidt operator satisfying the relation* $C_1 = C_2^{1/2}(I + S)\,C_2^{1/2}$.

Proof. The result is a consequence of [2] and Lemma 2. Indeed let $\bar{A} = \bar{C}_1^{-1} - \bar{C}_2^{-1}$ and $A = \hat{T}^{-1/2}(C_1^{-1} - C_2^{-1})\,\hat{T}^{-1/2}$. Then domain($A$) = range($\hat{T}^{1/2}C_2\hat{T}^{1/2}$) and domain($\bar{A}$) = range($\bar{C}_2$). To apply the lemma we must show that domain($\bar{A}$) = U(domain(A)). So let $y \in$ range($\bar{C}_2$). Then, for some x, $y = U\hat{T}^{1/2}C_2\hat{T}^{1/2}U'x$, so that U'(range($\bar{C}_2$)) $\subseteq$ range($\hat{T}^{1/2}C_2\hat{T}^{1/2}$). Conversely, if $y = \hat{T}^{1/2}C_2\hat{T}^{1/2}x$, then for $z = Ux$, $Uy = \bar{C}_2z$; that is, range($\hat{T}^{1/2}C_2\hat{T}^{1/2}$) $\subseteq U'$(range($\bar{C}_2$)). Thus, by Lemma 2, since $\bar{A}$ admits a closure $\bar{A}_c$, A admits a closure A_c and UA_cU' is closed. But UA_cU' extending UAU' must extend $\bar{A}_c$. We are going to show that actually $\bar{A}_c = UA_cU'$. First, we must check that both operators have the same domain. Let $x = UA_cU'y$, $x \in$ domain(UA_cU'). Choose $y_n \in$ domain(A), with $y_n \to U'y$ and $Ay_n \to$ some y_0. Then $Uy_n \to y$ and $UAU'(Uy_n) \to Uy_0$. Thus $y \in$ domain ($\bar{A}_c$) and $\bar{A}_cy = Uy_0$. But $A_cU'y = y_0$, so that $UA_cU'y = Uy_0$, or $\bar{A}_cy = UA_cU'y$. ∎

LEMMA 3. *Within the framework adopted,*

(a) *if* range(C_1) $\subseteq$ range(C_2), *then* range($\bar{C}_1$) $\subseteq$ range($\bar{C}_2$) *if and only if* $C_2^{-1}C_1$(range(E')) $\subseteq$ range(E');

(b) *if* range($\bar{C}_1$) $\subseteq$ range($\bar{C}_2$), *then* range(C_1) $\subseteq$ range(C_2) *if and only if* $(\bar{C}_2^{-1}\bar{C}_1)'$ (range(E)) $\subseteq$ range(E).

Proof. The assumption given in (a) is equivalent to $C_1 = C_2G$, G bounded linear on H. The one in (b) is equivalent to $\bar{C}_1 = \bar{C}_2\bar{F}$, $\bar{F}$ bounded linear on K.

We thus have, if G(range(E')) $\subseteq$ range(E'), $\bar{C}_1 = EC_2GE'$ and $GE' = E'\bar{G}$, $\bar{G}$ bounded linear on K. Consequently, range($\bar{C}_1$) $\subseteq$ range($\bar{C}_2$). Conversely, the latter inclusion gives $\bar{C}_1 = \bar{C}_2\bar{G}$, $\bar{G}$ bounded linear, so that $C_2GE' = C_2E'\bar{G}$, or G(range(E')) $\subseteq$ range(E').

The rest of the proof is quite similar. ∎

PROPOSITION 6. *Assume, within our framework, that* range(C_1) = range(C_2). *Then* range($\bar{C}_1$) = range($\bar{C}_2$) *if and only if* $C_2^{-1}C_1$(range(E')) $\subseteq$ range(E') *and* $C_1^{-1}C_2$(range(E')) $\subseteq$ range(E').

Proof. Apply Lemma 3 twice. ∎

LEMMA 4. *Assume that, within our framework,* S *and* C *commute* (see Sect. 2, Paragraph 2). *Then* T *is self-adjoint, commutes with* C, *and* $(T'T)^{1/2} = T$.

Proof. $TC^{1/2} = S$. But S and $C^{1/2}$ commute, so that $TC = SC^{1/2} = C^{1/2}TC^{1/2}$, which gives that T and $C^{1/2}$ commute. Since S is selfadjoint, $TC^{1/2} = C^{1/2}T'$, so that finally $T = T'$. ▮

PROPOSITION 7. *Assume that, within our framework, C_i and S_i commute for $i = 1, 2$. Then* range($\bar{C}_1$) = range($\bar{C}_2$) *if and only if* range(C_1) = range(C_2).

Proof. If range($\bar{C}_1$) = range($\bar{C}_2$), $\bar{C}_1 = \bar{C}_2\bar{G}$, $\bar{G}$ bounded linear, with bounded inverse. Thus $C_1TU' = C_2TU'\bar{G}$, or, by Lemma 4, $C_1 = C_2U'\bar{G}U'$. Now $U'\bar{G}U'$ has bounded inverse, so that range(C_1) = range(C_2). The converse is similar. ▮

Remark. When S_i can be manufactured so that S_i and C_i commute ($i = 1, 2$), the relationship between both operators can be more explicitely described as follows.

C_i and S_i commute if and only if $C_i = \sum_n a_n^{(i)} h_n^{(i)} \otimes h_n^{(i)}$, where the $a_n^{(i)}$'s are nonnegative and bounded, and the $h_n^{(i)}$'s form a family of eigenvectors of S_i.

When $\bar{C}_i$ has only eigenvalues $a_n^{(i)}$ of multiplicity 1, C_i and S_i commute if and only if $C_i = \sum_n b_n^{(i)} f_n^{(i)} \otimes f_n^{(i)}$, where the $b_n^{(i)}$'s are nonnegative and bounded, and $f_n^{(i)} = (a_n^{(i)})^{-1/2} C^{1/2} E' e_n^{(i)}$, with $e_n^{(i)}$ the eigenvector of $\bar{C}_i$ associated with $a_n^{(i)}$.

ACKNOWLEDGMENT

I wish to thank Professor C. R. Baker of the University of North Carolina at Chapel Hill for his help during the preparation of this work.

REFERENCES

1. A. F. GUALTIEROTTI, Extension of Gaussian cylinder set measures, *J. Math. Anal. Appl.* **57** (1977), 12–19.
2. C. R. RAO AND V. S. VARADARAJAN, Discrimination of Gaussian processes, *Sankhyā Ser. A* **25** (1963), 303–330.